

Was ist eine "gute" Umfrage?

Weichbold, Martin

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

Empfohlene Zitierung / Suggested Citation:

Weichbold, M. (2008). Was ist eine "gute" Umfrage? *SWS-Rundschau*, 48(3), 342-347. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-233346>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Methodik – Aktuell

Grenzen und Herausforderungen der Umfrageforschung

Was ist eine »gute« Umfrage?

Martin Weichbold (Salzburg)

Martin Weichbold: *Was ist eine »gute« Umfrage?* (S. 342–347)

Der Beitrag geht der Frage nach, wie die Qualität einer Umfrage beurteilt werden kann. Es gibt sehr unterschiedliche Ansätze, die sich in der Konzeption von Qualität und in der Umsetzung dieser Konzeption unterscheiden. In diesem Zusammenhang werden folgende Herangehensweisen erörtert: die klassischen Gütekriterien, die einer inhaltlichen Definition als Abweichung von einem wahren Wert folgen; *Codes of Ethics*, die sich auf die moralische Verantwortung der WissenschaftlerInnen beziehen; Standards bzw. Richtlinien, die den Forschungsprozess in viele Einzelschritte zergliedern; und die Qualitätssicherung mittels Zertifizierung von Sozialforschungsunternehmen, um Strukturen und Prozesse zu kontrollieren. Schließlich wird mit dem *Total Survey Error* ein integratives Konzept vorgestellt, das inhaltliche und prozessorientierte Qualitätskonzeptionen verbindet. Insgesamt zeigt sich, dass die Frage nach der Qualität einer Befragung sehr vielschichtig ist und sich nicht auf einen einzelnen Bewertungsmaßstab reduzieren lässt.

Schlagworte: Umfragen, Qualität, Qualitätssicherung

Martin Weichbold: *What is a »Good« Survey?* (pp. 342–347)

The article deals with the question, how to judge the quality of a survey. There are various approaches that differ considerably with regard to the conceptualisation and implementation of quality. The following concepts are discussed: the classical criteria of validity that follow a contentual definition as a deviation from a true value; *Codes of Ethics* that refer to the ethical responsibility of scientists; standards or guidelines, which structure the research process in a sequence of individual steps; and quality assurance that leverages certifications by social research organisations to control structures and processes. Finally, the *Total Survey Error* is being presented as an integrative approach that brings together content-based and process-orientated quality concepts. All in all it can be shown that the quality assessment of a survey represents a complex issue, which cannot be confined to a single rating scale.

Keywords: surveys, quality, quality assurance

1. Einleitung

Die Frage, wann eine Umfrage von guter Qualität ist, ist ebenso verständlich wie schwer zu beantworten. In Medien werden Berichte über empirische Sozialforschung oft mit dem Beiwort »repräsentativ« versehen, einem Begriff, dem in der breiten Öffentlichkeit offensichtlich eine andere Bedeutung beigemessen wird als in einem engeren Fachverständnis. Während der Begriff in der medialen Öffentlichkeit als generelles Qualitätslabel eingesetzt wird, das auf unverzerrte und somit richtige Ergebnisse verweisen soll, bezieht er sich streng genommen nur auf die strukturelle Übereinstimmung von Stichprobe und Auswahlgesamtheit (Schnell/ Hill/ Esser 2005, 304). Auch auf die Bekanntheit und den Ruf eines Sozialforschungsinstituts wird mitunter verwiesen, um damit die Qualität einer dort durchgeführten Studie zu belegen.

2. Klassische Gütekriterien

Sucht man im Bereich der Wissenschaft selbst nach Qualitätsmaßstäben, stößt man in der Literatur zunächst auf die klassischen Gütekriterien. Validität (Gültigkeit, also ob tatsächlich das Gesuchte erhoben wird), Reliabilität (Zuverlässigkeit, d. h. wie sehr mehrere Messungen ein übereinstimmendes Ergebnis bringen) und Objektivität finden sich nicht nur in Lehrbuchklassikern wie jenen von Friedrichs (1990, 14. Aufl.), Atteslander (2008, 12. Aufl.), Bortz/ Döring (2006, 4. Aufl.) oder Schnell/ Hill/ Esser (2005, 7. Aufl.), sondern auch in neu konzipierten Einführungen (Häder 2006, 1. Aufl., Weischer 2007). Sie beziehen sich darauf, wie gut ein Instrument in der Lage ist, den gesuchten wahren Wert zu messen, und bieten somit eine inhaltliche Konzeption von Qualität.

Auch wenn es umfangreiche und differenzierte Kritik an Anwendung, Anwendbarkeit und Annahmen der klassischen Gütekriterien gibt, sind diese seit Jahrzehnten fester Teil des Lehrkanons der empirischen Sozialforschung. Dies sagt freilich noch nichts über ihre Anwendung in der Praxis aus. Insbesondere im Bereich der kommerziellen Umfrageforschung sucht man Angaben zu den Gütekriterien oft vergeblich. Dies bedeutet allerdings nicht, dass in der angewandten Forschung keine Überlegungen zur Qualität von Umfragen angestellt werden – vor allem auf Ebene von Verbänden der Sozialforschungseinrichtungen sind dazu Anstrengungen in unterschiedlichen Dimensionen festzustellen.

3. Codes of Ethics

Berufsständische Vereinigungen und Verbände wie ADM, ESOMAR, AAPOR oder WAPOR¹ verfügen in der Regel über einen *Code of Ethics*, der allgemeine Regeln in Form von Anweisungen für (ethisch) richtiges Verhalten vorgibt. Der ESOMAR Internationale Kodex umfasst etwa acht Grundprinzipien und 14 zum Teil weiter differen-

1 ADM: Arbeitskreis deutscher Markt- und Sozialforschungsinstitute, www.adm-ev.de ; ESOMAR: European Society for Opinion and Marketing Research, www.esomar.nl/ ; AAPOR: American Association for Public Opinion Research, www.aapor.org/ ; WAPOR: World Association for Public Opinion Research, www.unl.edu/WAPOR/ .

zierte Artikel. Die Abstraktionsebene ist somit notwendigerweise hoch. Neben Grundprinzipien wie Redlichkeit und berufsständischer Verantwortung wird auch auf Datenschutz, Umgang mit Befragten wie Auftraggebern oder auf die Offenlegung der eingesetzten Verfahren eingegangen.

Bemerkenswert an dieser Form der Qualitätssicherung ist, dass Moral und Verantwortungsbewusstsein der einzelnen ForscherInnen eine hohe Qualität der Untersuchung sicherstellen sollen, Qualitätssicherung hier also über ethische Verhaltensanweisungen der Beteiligten erfolgt. Was als qualitativ gilt, ist Angelegenheit der professionellen Einschätzung des bzw. der Einzelnen.

4. Richtlinien und Standards

Aufgrund ihrer knappen Formulierung können *Codes of Ethics* nicht im Detail auf konkrete Designs, Forschungsschritte oder Problemlagen eingehen. Ihnen sind Richtlinien oder Standards beigegeben, die Minimalanforderungen für die Durchführung von Forschungsprojekten beschreiben. In den unterschiedlichen Verbänden ist dies nicht einheitlich gelöst: ESOMAR hat mehrere Richtlinien (*guidelines*) zu unterschiedlichen Bereichen herausgegeben, ADM verfügt mit den »Standards zur Qualitätssicherung in der Markt- und Sozialforschung« (ADM 1999) über ein mit rund 60 Seiten ausführliches und umfassendes Dokument.

Sechs Hauptkapitel sind in bis zu sieben Unterpunkte gegliedert, die wiederum weiter unterteilt sind. Dadurch entsteht eine stark strukturierte Aufgliederung der Thematik, die Festlegungen für den Forschungsablauf, aber kaum konkrete inhaltliche Festlegungen trifft. Für die »Größe und Qualität der Stichprobe« wird etwa festgehalten, dass sie »sich in erster Linie an der Problemstellung der Untersuchung [orientieren]. Bei qualitativen Stichproben muss aufgrund der geringen Fallzahlen besondere Sorgfalt bei der Selektion der teilnehmenden Personen aufgewendet werden. Die Größe der Stichprobe ist so zu bemessen, dass ... eine hinsichtlich des jeweiligen Forschungsziels vertretbare Fehlertoleranz vorliegt« (ADM 1999, Kap. IV/ A/ 2., 30).

Die einzelnen ForscherInnen sind also (noch) nicht aus ihrer moralischen Verantwortung entlassen – durch die fein strukturierte Aufzählung der verschiedenen Bereiche bieten die ADM-Standards aber einen Katalog relevanter Themen und potenzieller Problemfelder für eine Qualitätsbewertung und -sicherung. Qualität wird über einzelne Ablaufschritte, nicht über inhaltliche Definitionen festgelegt.

5. Zertifizierung

Die ADM-Standards sind insofern nahe an einem weiteren Schritt zur Qualitätssicherung, der Zertifizierung nach ISO². Dabei kann sich ein Unternehmen die Implemen-

2 Kurzbezeichnung für die Internationale Organisation für Normung. Ausgehend vom ursprünglichen Bestreben, technische Produkte zu normieren, um einen leichteren internationalen Austausch zu ermöglichen, wurden zunehmend (Herstellungs-) Verfahren und schließlich auch Dienstleistungen standardisiert.

tierung eines Qualitätsmanagementsystems von einem autorisierten Zertifizierungsunternehmen bestätigen lassen, was Wettbewerbsvorteile schaffen soll. Während bisher eine Zertifizierung nur nach den allgemeinen Kriterien für Qualitätssicherung (ISO-Norm 9000ff.) oder einzelnen nationalen Normen möglich war und sie von Sozialforschungsunternehmen kaum genutzt wurde, steht seit 2006 mit der ISO-Norm 20252 eine spezifisch für die Belange der Markt-, Meinungs- und Sozialforschung konzipierte Regelung zur Verfügung (Wiegand 2007).

Wer in der ISO-Norm eine inhaltliche Festlegung der Qualität etwa von Umfragen sucht, wird enttäuscht. Qualität soll über zwei Dimensionen erreicht werden: zum einen über organisatorische Rahmenbedingungen (etwa, dass die Dokumentation eines Forschungsprojekts bestimmte Anforderungen erfüllen muss), zum anderen über Verfahrensvorschriften (z. B. Anweisungen, wie auf eine Anfrage zu reagieren ist oder welche Informationen ein Untersuchungszeitplan zu enthalten hat). Die Einhaltung der »Norm« erfolgt durch Kontrolle (und bleibt nicht mehr nur der individuellen ethischen Verantwortung überlassen), was eine weitgehende Dokumentation aller Schritte voraussetzt. Offen bleibt, ob durch die Struktur- und Prozesskontrolle ein (inhaltlich) qualitätsvolles Produkt garantiert wird: Wie lässt sich also die inhaltliche Dimension wieder hereinholen?

6. Total Survey Error

Seit einigen Jahren wird in der englischsprachigen Methodenliteratur ein Ansatz verfolgt, der im deutschen Sprachraum erst sehr zaghaft wahrgenommen wird. Der Ansatz des *Total Survey Error* (TSE) (Biemer/ Lyberg 2003, Weisberg 2005) verbindet Prozesssicht und inhaltliche Überlegungen zur Qualität einer Umfrage wieder miteinander und integriert beides in einem Gesamtkonzept. Grundsätzlich wird Qualität als multidimensionales Konstrukt aufgefasst, das allgemein als *fitness for use* (Brauchbarkeit im Hinblick auf den beabsichtigten Verwendungszweck) definiert werden kann. Je nach Anforderungen lassen sich zahlreiche Qualitätskriterien ableiten, von denen lediglich eines, nämlich *accuracy* (Exaktheit) mit dem TSE erfasst wird. Andere Qualitätsmerkmale, etwa Relevanz, Kohärenz oder Verfügbarkeit, werden in dem Konzept nicht berücksichtigt.

Das Konzept basiert auf dem Axiom der klassischen Testtheorie, wonach sich der gemessene Wert aus wahren Wert und Fehler zusammensetzt, wobei dieser Fehler auf vielfältige Einflüsse zurückzuführen ist. Dieser Gesamtfehler besteht aus dem *sampling error* (dem bekannten Stichprobenfehler) und dem *non-sampling error* – das sind alle anderen Fehler im Rahmen der einzelnen Phasen und Entscheidungen des Forschungsprozesses.

Der TSE bezieht sich auf eine inhaltliche Vorstellung von Qualität: Ähnlich wie bei den Gütekriterien Validität und Reliabilität geht es um die Abweichung von einem wahren Wert – damit unterstellt man auch die Existenz eines solchen. Dabei erfolgt Qualitätskontrolle mit Hilfe von Prozesskontrolle, indem eine Umfrage (*survey*) als eine Abfolge aufeinander bezogener Entscheidungen begriffen wird. In dieser Hinsicht

ist der TSE ein integratives Konzept, das inhaltliche wie prozessuale Aspekte kombiniert.

7. Resümee

Nicht nur »Qualität« ist ein multidimensionales Konstrukt, sondern offenbar auch die Bestimmung derselben. Es wurden verschiedene Ansätze gezeigt, die sich deutlich unterscheiden, aber auch Verbindungspunkte aufweisen.

Tabelle 1: Qualitätsdefinitionen und -beurteilungen im Überblick

Bezeichnung	Definition von Qualität	Art der Qualitätsbeurteilung/ -kontrolle
Klassische Gütekriterien	Inhaltlich bestimmt (Abweichung vom wahren Wert)	Kriterien, statistische Koeffizienten
<i>Codes of Ethics</i>	Keine nähere inhaltliche Definition, professionelle Einschätzung der einzelnen WissenschaftlerInnen	Ethisches Verhalten Einzelner
Standards/Richtlinien	Katalogartige Zergliederung in viele kleine Teilbereiche, aber keine inhaltliche Definition	Appell an Sorgfalt bei Durchführung, Kontrolle
Zertifizierung	Anforderungen an Strukturen und Prozesse, keine inhaltliche Definition	Dokumentation und Kontrolle
<i>Total Survey Error</i>	Inhaltliche Definition (Abweichung vom wahren Wert) und prozessuale Orientierung	Analytische oder quantitative Bestimmung des Fehlers

Quelle: Eigene Zusammenstellung

Eine Leitdifferenz bei der Bestimmung von Qualität ist jene zwischen Inhalt und Form. Einige Ansätze versuchen eine inhaltliche Konzeption von Qualität, andere verzichten von vornherein auf eine inhaltliche Qualitätsdefinition und beschränken sich auf die Kontrolle von Strukturen und Prozessen. Dazwischen sind jene Ansätze einzuordnen, die sich zwar von inhaltlich definierten Qualitätsmaßstäben leiten lassen, diese aber nicht selbst spezifizieren, sondern dies der *professionellen Einschätzung* und ethischen Verantwortung des bzw. der Einzelnen überlassen.

Dem zugrunde gelegten Qualitätskonzept folgen auch die Wege der Qualitätskontrolle. Bei inhaltlich festgelegten Kriterien liegt es nahe, diese auch zu messen bzw. zu bestimmen. Bemerkenswert ist, dass der moralische Appell an eine gewissenhafte Durchführung nicht nur in den *Codes of Ethics* zu finden ist, sondern auch noch in Standards und Richtlinien. Im nächsten Schritt, bei der Zertifizierung von Sozialforschungsinstituten, will man sich aber offensichtlich nicht mehr auf die Integrität der ForscherInnen verlassen: An ihre Stelle treten weit reichende Dokumentations- und Kontrollpflichten.

So pendelt die Qualitätsdiskussion zwischen Inhalten und Prozessen, zwischen Vertrauen und Kontrolle. Es wird klar, dass es »die« Qualität nicht gibt: So wie »gute« Sozialforschung aus einem komplexen Prozess mehrdimensionaler, aufeinander bezogener Handlungen und Entscheidungen entsteht, wird man sie letztlich nicht mit einer einzigen Kennzahl, einem wohlklingenden Begriff oder einem Gütezeichen umfassend bewerten können.

Literatur

- ADM (Arbeitskreis deutscher Markt- und Sozialforschungsinstitute) (1999) *Standards zur Qualitätssicherung in der Markt- und Sozialforschung*, verfügbar unter: www.adm-ev.de/pdf/quali.pdf, 26. 6. 2008.
- Atteslander, Peter (2008) *Methoden der empirischen Sozialforschung*. Berlin (12. Auflage).
- Biemer, Paul P./ Lyberg, Lars E. (2003) *Introduction to Survey Quality*. Hoboken.
- Bortz, Jürgen/ Döring, Nicola (2006) *Forschungsmethoden und Evaluation*. Berlin u. a. (4. Auflage).
- DIN (Deutsches Institut für Normung) (2006) *DIN ISO 20252. Markt-, Meinungs- und Sozialforschung. Begriffe und Dienstleistungsanforderungen*. Berlin.
- ESOMAR (European Society for Opinion and Marketing Research) (2007) *ICC/ESOMAR Internationaler Kodex für die Markt- und Sozialforschung*, verfügbar unter: <http://www.esomar.org/index.php/codes-guidelines.html>, 26. 6. 2008.
- Friedrichs, Jürgen (1990) *Methoden empirischer Sozialforschung*. Opladen (14. Auflage).
- Häder, Michael (2006) *Empirische Sozialforschung. Eine Einführung*. Wiesbaden.
- Schnell, Rainer/ Hill, Paul B./ Esser, Elke (2005) *Methoden der empirischen Sozialforschung*. München/ Wien (7. Auflage).
- Weisberg, Herbert F. (2005) *The Total Survey Error Approach. A Guide to the New Science of Survey Research*. Chicago.
- Weischer, Christoph (2007) *Sozialforschung*. Konstanz.
- Wiegand, Erich (2007) *Standards und Normen in der Markt- und Sozialforschung*. In: König, Christian u. a. (Hg.) *Qualitätsmanagement und Qualitätssicherung. GESIS-Tagungsberichte*, Bd. 1. Bonn, 43–50.
- Kontakt:*
martin.weichbold@sbg.ac.at